

Ecological Diagnostics for Marine Mammals: Appraisal
of molecular-based methods for dietary and age
estimation

Glenn John Dunshea

BSc (Hons) GCRC

Submitted in fulfillment of the requirements for the degree of
Doctor of Philosophy

Institute of Marine and Antarctic Studies

University of Tasmania

April 2012

“The ballast of factual information, so far from being just about to sink us, is growing less daily. The factual burden of a science varies inversely with its degree of maturity”

Sir Peter Medawar, OM. “Two conceptions of science”,
Henry Tizard Memorial Lecture, Encounter 143, August 1965

Declarations

Originality;

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of the my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Signed:

Dated:

Glenn J Dunshea

Access;

This thesis may be made available for loan. Copying and communication of any part of this thesis is prohibited for two years from the date this statement was signed; after that time limited copying and communication is permitted in accordance with the Copyright Act 1968.

Signed:

Dated:

Glenn J Dunshea

Ethics;

The research associated with this thesis abides by the international and Australian codes on human and animal experimentation, the guidelines by the Australian Government's Office of the Gene Technology Regulator and the rulings of the Safety, Ethics and Institutional Biosafety Committees of the University.

This research was approved by the Animal Ethics Committee of the University of Tasmania (Permit # A0008315).

Signed:

Dated:

Glenn J Dunshea

Peer reviewed literature already published from parts of this thesis

Paper 1:

Dunshea, G., N. B. Barros, R. S. Wells, N. J. Gales, M. Hindell, A. and S. N. Jarman (2008)
Pseudogenes and DNA-based diet analyses; a cautionary tale from a relatively well sampled predator-prey system. *Bulletin of Entomological Research* 98: 239-248.

Paper 2:

Dunshea, G. (2009). DNA-based diet analysis for any predator. *PLoS One* 4(4): e5252.doi:10.1371/journal.pone.0005252.

Paper 3:

Garde E, A.K. Frie, **G Dunshea** , SH Hansen, KM Kovacs, C Lydersen (2010) Harp seal ageing techniques-teeth, aspartic acid racemization, and telomere sequence analysis. *Journal of Mammalogy*, 91: 1365–1374.

Paper 4:

Dunshea, G., D. Duffield, N. Gales, M. Hindell, R. S. Wells and S. N. Jarman (2011).
Telomeres as age markers in vertebrate molecular ecology. *Molecular Ecology Resources*, 11: 225-235.

Statement of Co-Authorship

The following people and institutions contributed to the publication of the work undertaken as part of this thesis:

- Paper 1 **G. J. Dunshea** (70%), *N. B. Barros* (5%), *R. S. Wells* (5%) *N. J. Gales* (5%)
M. A. Hindell (5%), *S. N. Jarman* (10%)
- Paper 3 **G. J. Dunshea** (20%), *Garde E* (35%), *A.K. Frie* (25%), *S.H. Hansen* (10%),
K.M. Kovacs (5%), *C Lydersen* (5%)
- Paper 4 **G. J. Dunshea** (65%), *D. Duffield* (5%), *N. Gales* (5%), *M. Hindell* (10%),
R. S. Wells (5%), *S. N. Jarman* (10%)

Details of the Authors roles:

Paper 1

G. J. Dunshea contributed to the idea of the manuscript, its formalisation and development, performed the laboratory work, data collection, data analysis and wrote the paper

S. N. Jarman contributed to the idea of the manuscript, its formalisation, development and manuscript refinement and presentation

N. B. Barros and R. S. Wells assisted with sample collection and manuscript refinement and presentation

N. J. Gales and M. A. Hindell assisted with manuscript refinement and presentation

Paper 3

G. J. Dunshea, contributed to the idea of the paper, provided harp seal telomere assay data, data analysis and interpretation, wrote the telomere part of the paper and assisted with manuscript refinement and presentation

Garde E. contributed to the idea of the paper, its formalisation and development , provided aspartic acid racemization data, its analysis and interpretation and wrote the bulk of the paper

A.K. Frie and S.H. Hansen contributed to the idea of the paper, its formalisation and development , provided teeth growth layer group analysis, sample collection and manuscript refinement and presentation

K.M. Kovacs and C Lydersen contributed to the idea of the paper, its formalisation and development, performed data analysis and manuscript refinement and presentation

Paper 4

G. J. Dunshea, contributed to the ideas in the paper, their formalisation and development, performed most of the laboratory work, data collection, data analysis and wrote the paper

D. Duffield assisted with providing extracted DNA from historically archived samples

N. Gales contributed to the ideas in the paper, their formalisation, development and manuscript refinement and presentation

M. Hindell contributed to the ideas in the paper, their formalisation, development and manuscript refinement and presentation

R. S. Wells assisted with sample collection and manuscript refinement and presentation

S. N. Jarman contributed to the ideas in the paper, their formalisation, development and manuscript refinement and presentation

We the undersigned agree with the above stated “proportion of work undertaken” for each of the above published (or submitted) peer-reviewed manuscripts contributing to this thesis:

Signed: _____

Mark Hindell

Mike Coffin

Supervisor

Director

Institute of Marine and Antarctic Studies
Studies

Institute of Marine and Antarctic

University of Tasmania

University of Tasmania

Date: _____

Table of Contents

Acknowledgements.....	1
Abstract.....	2
Chapter 1 - Introduction	7
1.1 Diet Investigations	7
1.2 Age-based Investigations	10
1.3 Structure of the thesis	13
Chapter 2 - A universal DNA-based diet technique for application to potentially any predator ..	15
2.1 Introduction	15
2.2 Materials and Methods.....	16
2.2.1 In silico development	16
2.2.2 Development of laboratory protocol.....	19
2.2.3 Data analysis	22
2.3 Results	23
2.3.1 Taxon specific presence/absence of Pac I restriction site in amplicon.....	23
2.3.2 'Universal' Complementarity of PCR primers	24
2.3.3 Assessment of target 16SPLSU fragment to identify species from sequence data	27
2.3.4 SSCP scoring of identical clones	30
2.3.5 Removal of predator amplicons.....	30
2.3.6 Prey detection from captive dolphin scat.....	31
2.3.7 Prey detection from free ranging Sarasota Bay dolphin scat	32
2.4 Discussion.....	34
Chapter 3 - DNA-based diet analyses to describe cetacean diet; experimental and empirical evaluations.....	39
3.1 Introduction	39
3.2 Methods	40
3.2.1 Initial Tests of Efficacy of Different Protocols for Dolphin faeces DNA Extraction.....	40
3.2.2 Contamination Control of DNA Extracts and PCR Assays	41
3.2.3 Samples Used	41
3.2.4 Universal Assays with SWBD samples.....	43
3.2.5 PCR Primer Design and Testing.....	44
3.2.6 Real Time PCR screening for Mugil cephalus DNA in SWBD samples	45

3.2.7 Universal Assays with Sarasota Bay samples	45
3.2.8 Data Analysis	45
3.3 Results	48
3.3.1 DNA Extraction Test	48
3.3.2 Contamination Control of DNA Extracts and PCR Assays.....	48
3.3.3 SWBD Universal Assays	49
3.3.4 SWBD Mugil cephalus DNA assays	54
3.4 Sarasota Bay dolphin Universal Assays	56
3.4.1 Prey Assignment by BLAST threshold and Placement in Neighbour Joining Analysis...	57
3.4.2 Prey Assignment Using the MOTU/Statistical Assignment Package method	59
3.4.3 Comparison of different prey taxonomic assignment techniques.....	59
3.4.4 Occurrence and Inferred Relative Importance of Sarasota Bay dolphin prey	64
3.4.5 Comparison Sarasota Bay dolphin diet estimated via DNA and stomach contents data	64
3.5 Discussion.....	68
3.5.1 Captive Feeding trials and SWBD prey detection results.....	68
3.5.2 Sarasota bay data	70
3.5.3 Conclusion	73
Chapter 4 - Pseudogenes and DNA-based diet analyses: a cautionary tale from a relatively well sampled predator-prey system	75
4.1 Introduction.....	75
4.2 Materials and Methods	77
4.2.1 Sample collection and analysis.....	77
4.2.2 Sequence scoring and analysis.....	77
4.2.3 Confirmation of NUMT origin of spurious sequences.....	78
4.3 Results	78
4.3.1 pNUMT frequencies, proportions and PCR characteristics.....	78
4.3.2 NUMT sequence characteristics, phylogenetic analysis and substitution pattern	81
4.3.3 Confirmation of NUMT origin of sequences	81
4.4 Discussion.....	85
Chapter 5 – Telomeres as age markers for marine mammals	89
5.1 Introduction.....	89

5.2 Methods	91
5.2.1 General Characteristics of Study Populations and Sample Sources	91
5.2.2 Bottlenose Dolphin samples	92
5.2.3 Southern Right Whales	95
5.2.4 Harp Seal Samples.....	95
5.2.5 Data Analysis	97
5.2.6 qPCR Assay Parameter Information.....	98
5.3 Results	99
5.3.1 Bottlenose Dolphins.....	99
5.3.2 Southern Right Whale qPCR assays	105
5.3.3 Harp Seals.....	105
5.4 Discussion.....	108
5.4.1 Bottlenose Dolphins and Southern Right Whales.....	108
5.4.2 Harp Seals.....	110
5.4.3 Methodological Considerations	112
Chapter 6 – General Discussion	113
6.1 DNA-based methods to study cetacean diet	113
6.2 Live marine mammal age estimation and telomeres	120
6.2.1 Telomeres for age estimation: Reappraisal of the literature	120
6.3 Conclusions	128
References	129

List of Figures

Figure 2.1 Proportion of species in alignments provided at <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0005252#s5> (number of species in each alignment provided in parentheses – x-axis) that have priming regions directly complimentary at the first 8- 14 nucleotide positions of the 3' ends to the primers presented in the main manuscript (grey bars). Hatched bars represent the same analysis when the additional primers (5'-3'): AAGACCCCGTTGAGCTT, AAGACCCTGTCGAGCTT, AAGACCCTATCGAGCTT, AAGACCCTTTGGAGCTT, AAGACCCTATAAACTT, AAGACCCTGTGGAAGCTT, AAGACCCTATCGAAGCTT, AAGACCCTATAAATCTT and AAGACCCTATAGATCTT are included. Exact proportion is indicated above bars..... 25

Figure 2.2 Test of universal primers designed for the small fragment of 16s mtDNA targeted in this study. Lanes: M : 1kb ladder; 1: Mammalia: *Arctocephalus pusillus doriferus*, 2: Mammalia: *Pagophilus groenlandicus*, 3: Mammalia: *Tursiops truncatus*, 4: Aves: *Aptenodytes patagonicus*, 5: Aves: *Stercorarius* sp. 6: Teleosti: *Sardinops sagax*, 7: Teleosti: *Trachurus novaezelandiae*, 8: Echinodermata: *Centrostephanus rodgersii*, 9: Echinodermata: Crinoidea sp. 10: Mollusca: *Nototodarus* sp. 11: Mollusca: *Nototodarus gouldi* 12: Mollusca: Octopoda sp. 13: Mollusca: Pteriomorpha sp. 14: Crustacea: Amphipoda sp. 15: Crustacea: Penaeidae sp. 16: Crustacea: *Thysanoessa macrura*, 17 Insecta: Lepidoptera sp. 1, 18: Insecta: Lepidoptera sp. 2. 19: PCR no template control..... 27

Figure 2.3 Comparison of vertebrate (top graph) and invertebrate (bottom graph) within species K2P distance (x-axis) and between species K2P distance (y-axis) in taxa from Table 2.3. Diagonal represents equal within and between species K2P distances. Points falling on or below diagonal line represent species that have equal or higher intra-specific divergence compared to the pairwise divergence with at least one other species. In such cases specific identification from sequence data based on distance measures would be ambiguous..... 29

Figure 2.4. Example of SSCP gel used to identify identical clones from a clone library. M = molecular size marker, NTC = No template control of PCR, E3, F3... = Different clones by well position in 96 well plate. V1, V2... = Variant clones (i.e. All V1 clones are identical). Clone identities confirmed by sequencing: V1: *Sardinella lemuru/Sardinops sagax*, V2: *Scomber australasicus*, V3: *Trachurus novaezelandiae*, V4: Chimera of *Trachurus novaezelandiae* and *Sardinella lemuru/Sardinops sagax* sequences, V5: *Tursiops truncatus*, V6: *Sardinella lemuru/Sardinops sagax*, 2 substitutions difference from V1, V7: *Arripis georgianus*, V8: *Scomber australasicus* 1 substitution difference from V2, V9: Poor sequence – discarded, V10: *Sillago robusta* 30

Figure 3.1 Frequency histogram of approximate SWBD scat volumes. The numbers of final samples used differ from the totals since the original DNA extract may have been cross-contaminated during DNA extraction procedures and there was no sample remaining to perform a second DNA extraction 42

Figure 3.2 Priming regions of *Mugil cephalus* species specific primers designed in this study. Dots represent same nucleotide as *Mugil cephalus* for that nucleotide position, ~ represents the gap between priming regions. Note the mismatch between the PCR primers (particularly at the 3' end of the primers) for non-target species. 44

Figure 3.3 Bland-Altman plot of PCR threshold cycle (c(t)) differences between DNA extraction kits. Bold middle line represents the mean difference between MoBio and QIAGEN kits (MoBio c(t) – QIAGEN c(t)) and the dotted line the standard deviation of the mean. 48

Figure 3.4 Proportion of predator origin (mtDNA and NuMT) clones from different sample volume categories of SWBD samples. Sample volume categories and number assayed indicated on the x-axis. Filled circles are individual data points, the open circles the mean of each volume category, whiskers bootstrapped 95% binomial confidence interval for the mean for grouped data. 51

Figure 3.5 Actual Proportion of prey in SWBD diet (Dark Grey horizontal lines) compared to proportions of amplicons in clone libraries calculated by different means. 1) Prey proportions averaged across individual scat samples (circles), 2) Prey proportions averaged across individual dolphins (squares), and 3) absolute total proportion (triangles) by pooling all prey amplicons in clone libraries. Whiskers represent the bootstrap 2.5 and 97.5 percentiles for 95% confidence intervals (light grey lines) and 95% confidence intervals for the mean (red lines) due to multinomial sampling error. Symbols represent arithmetic mean between groups (circles and squares) or total prey proportion (triangle). For comparison, arithmetic 95% confidence intervals of arithmetic mean proportions of raw data are given (dotted lines). Prey species as follows: Sa: *Scomber australasicus*, Ag: *Arripis georgianus*, Mc: *Mugil cephalus*, N: *Nototodarus spp.*, P: *Penaeidae spp.* 53

Figure 3.6 Samples that tested positive for *Mugil cephalus* DNA per day of the feeding trial (sequential days from left to right). For the first day, only samples collected after the first positive sample were included in this figure. Excluding the first test day, numbers above x-axis are total sample numbers for that day. Data are black dots with actual value to the right of the data point. Data are joined by lowess smoothed line. 55

Figure 3.7 Boxplot with overlaid probability density function estimated with kernel density estimates of proportion of positive samples from bootstrap replicates ($n = 1000$) from first to last prey detection. Filled red line is the actual proportion of positive samples over the

duration of the prey detection period; dotted red lines are the bootstrap 95% confidence intervals of the proportion of positive samples calculated from the bootstrap variance estimate. 66

Figure 3.8 Average taxon (top panel: species; bottom panel: family) amplicons per clone library. Filled circles represent arithmetic average across individuals, open circles represent proportion of pooled total prey amplicons. Whiskers are 95% confidence intervals for 1) population arithmetic mean (dotted lines of filled circles); 2) bootstrapped 95% confidence intervals for mean due to sampling error in each clone library (red line of filled circles) and 3) bootstrapped 95% confidence intervals for proportion of pooled total prey amplicons due to sampling error in each clone library (filled lines of open circles) 66

Figure 3.9 Absolute frequency of occurrence (sensu Wright, 2010) of six most common prey species in Barros & Wells (1998) compared to this study. Whiskers represent exact 95% binomial confidence intervals. 67

Figure 3.10 Average numerical proportion of prey items across all samples (other studies) and average proportion of prey item clones across all sample (this study). Whiskers represent truncated 95% confidence intervals of arithmetic mean. Variance was not available for Berens et al. (2010). 67

Figure 4.1 Characteristics of putative NUMTs recovered from all samples **(A)** Frequency histogram of the proportion of putative NUMTs in the clone library from each sample ($n = 15$); first category from 0-0.1 includes libraries with no pNUMTs. **(B)** Relationship between the proportion of putative NUMTs in the library (x-axis) and the diversity of prey species identifiable (filled circles) and putative NUMT haplotypes (open squares) (y-axis). These relationships were not statistically tested as the variables are not independent. **(C)** Relationship between the threshold PCR cycle of the 16S PCR used to amplify prey DNA (x-axis) and the diversity of the prey species identified (y-axis) (Kendall Tau correlation; $\tau = -0.45$, $z = -2.24$, $p = 0.03$) **(D)** Relationship between the threshold PCR cycle of the 16S PCR used to amplify prey DNA (x-axis) and the proportion of putative NUMTs in libraries (y-axis) (Kendall Tau correlation; $\tau = 0.44$, $z = 2.29$, $p = 0.02$) 80

Figure 4.2 Minimum Evolution phylogenetic tree displaying the relationship between putative NUMT haplotypes and cetacean mtDNA. Other laurasiatherian mtDNA was used for an outgroup (bottom seven branches). Topology was tested by bootstrapping with 1000 replications and the consensus tree is shown. Only values at nodes with a bootstrap score of >50% are shown. Note the grouping of all putative NUMT haplotypes within the major cetacean clade in relation to the outgroup but still distal to and highly diverged from the majority of cetaceans. GenBank accession numbers are displayed after species name..... 84

Figure 5.1 Initial terminal restriction fragment experiment on *Tursiops truncatus* DNA from multiple individuals. Different individuals indicated by sample identifier at top of image. Samples with * indicate samples of the same individual collected different years. Lines indicate problems with sample seepage from bottom of gel well. Dark lines in outer lanes represent Lambda *Hind* III size markers (top marker concatenation of 23130bp & 4361bp fragments (27691bp)) with lines drawn between the same size markers on either side of gel to account for uneven fragment migration and/or blotting distortion across gel. 99

Figure 5.2 *Bal* 31 time series digestion of DNA from a single individual to characterize telomere sequence distribution in bottlenose dolphins. **1.** Gel image of experiment. DNA size in base pairs indicated down left side. Lanes; **1-7** represent 1/6 volume of total *Bal* 31 treatment - 0 (control), 10, 20, 30, 40, 60 and 120 minutes of *Bal* 31 digestion. Lanes **8-14** represent remaining 5/6 volume from lanes **1-7** further treated with restriction enzymes as per a conventional terminal restriction fragment assay. **2.** Density plot relative copy number in each DNA size class from restriction digestion treated DNA from the same experiment. Digest times represented in top left of graph. Note the rapid digestion of DNA in the ~23-4 kb portion after 60 minutes of digestion; this fraction represents true telomeric DNA (Represented by **a** in both panels). Note also the resistance to *Bal* 31 digestion of laddering in smaller size classes and of DNA up to ~8.5kb that overlaps with true telomeric DNA (represented by **b** in both panels). 100

Figure 5.3 Boxplot of telomere sequence per genome values from bottlenose dolphins. Bottom whisker is 25th percentile, box is 25th-75th percentile with median dark line, top whisker above 75th percentile. Outliers (open circles) are defined as over 1.5x the interquartile range. Note the two extreme outliers, both from historical samples from 1995. 101

Figure 5.4 Telomere sequence per genome values of longitudinal samples from bottlenose dolphin individuals collected at different ages. Whiskers for data points represent telomere sequence per genome value standard deviation..... 102

Figure 5.5 Cross sectional telomere sequence per genome values for Sarasota Bay bottlenose dolphin samples. Where two data points are from the same individual refer to the legend on the upper right. There was no significant relationship between telomere sequence per genome and age with this data..... 103

Figure 5.6 Boxplot of telomere values obtained by densitometric analysis of initial TRF assays performed in Fig. 5.1. Bottom whisker is 25th percentile, box is 25th-75th percentile with median dark line, top whisker above 75th percentile. Note the lack of variation across values. 104

Figure 5.7 Densitometric profiles from representative lanes from Fig. 5.1. Legend show representative lane from samples, age of samples in parentheses. FB19 (49 years old) had the highest telomere metric values and FB157 (An adult of unknown age) had the lowest telomere metric value. 104

Figure 5.8 Telomere sequence per genome values from 12 samples from 6 cow-calf pairs. Cows are grey bars, calfs white bars, whiskers are standard error. 105

Figure 5.9 Example of Harp Seal TRF autoradiograph. 106

Figure 5.10 Plot of harp seal telomere sequence per genome estimates and tooth age. There was no significant relationship between the two variables. Whiskers represent standard deviation of CV-filtered telomere metrics. 107

Figure 5.11 Plot of harp seal telomere metrics gained from densitometric analysis of TRF assays and tooth age. There was no significant relationship between the variables with all data, however if the two points from 1 year old males are excluded there is a significant relationship between the variables. There was no valid reason to exclude these data points. 107

List of Tables

Table 2.1 Proportions of species that contain the <i>Pac</i> I restriction site within the amplicon region out of species represented on GenBank (as of September 2006).....	24
Table 2.2 Forward primers tested for 3' complementarity with sequences provided in supplement 2. Primers 1-10 are present in the degenerate forward primer presented in the main manuscript and used in the study and primer pair empirical trial. Primers 11-19 are not represented within the degenerate forward primer in the main manuscript; nucleotide positions where these primers differ from those used in the study are shown in italicised, bold and larger font.	26
Table 2.3 Summary of comparisons of within species divergence with between species divergence of species within each familial (or closest higher taxon) group for 16S mtDNA target fragment.....	28
Table 2.4 Prey and mammal clones identified in clone libraries from 5 captive feeding trial samples subject to Treatment 2 (<i>Pac</i> I digestion for predator DNA exclusion prior to PCR and also prior to cloning). Prey species abbreviations are shown below.). Frequency of prey occurrence across scats is presented in the 'FOC' column.	32
Table 2.5 Prey operational taxonomic units (OTUs) identified from scat samples obtained from free-ranging Sarasota Bay <i>T. truncatus</i> (n = 5 different individuals). Frequency of prey occurrence across scats is presented in the 'FOC' column.	33
Table 3.1. Summary of SWBD scat samples available for assays after DNA extraction and contamination tests. Numbers in parentheses are the actual number of samples collected in that day. Difference between samples used and numbers in parentheses are due to material being unavailable for a second round of DNA extraction following contamination tests.....	41
Table 3.2 Summary of SWBD scat samples available for assays after DNA extraction and contamination tests	43
Table 3.3 Raw data from universal prey detection assays of 12 SWBD samples. Samples with ✓ indicate this sample was assayed as part of the first chapter and the data are included here. Dashed lines indicate different days of the feeding trial.	50
Table 3.4 Nuclear mitochondrial pseudogenes (NuMts) variants isolated from SWBD clone libraries	51

Table 3.5 Frequency of Occurrence (FOC) scores weighted by proportion of total prey clones for SWBD samples. Prey groups arranged in order from most to least important according to weighted FOC scores.....	52
Table 3.6 Summary of OTUs identified from Sarasota Bay <i>T. truncatus</i> faeces and gastric samples using a Threshold/Topological Neighbour Joining Analysis.....	58
Table 3.7 Summary of Sarasota Bay <i>T. truncatus</i> number of clones per individual taxonomically assigned to prey MOTUs using SAP software and accounting for <i>Taq</i> DNA polymerase error in clone libraries. If Genus and Family level identifications were > 95% similar to another identified taxon, they are listed below the most similar (% identity) species level identification and indented. Blue text indicated a difference in prey assignment between the '% Threshold/Neighbour joining' method and the MOTU/SAP method.....	62
Table 3.8 Species and family Frequency of Occurrence (FOC) and FOC scores for taxa appearing in more than one sample. Taxa are grouped in descending order of weighted frequency of occurrence score.....	65
Table 4.1 Summary of the occurrence between samples of all recovered putative NUMT haplotypes and their BLAST closest matches.....	79
Table 4.2 Number of pairwise nucleotide differences (bottom diagonal) and pairwise genetic distances as estimated by the Kimura 2 parameter substitution model (top diagonal) between suspected NUMT sequences, <i>T. truncatus</i> and the closest BLAST match of the suspected NUMT's. Light grey shaded areas are the pairwise comparisons between suspected NUMTs and true cetacean mtDNA, dark grey shaded areas are the lowest genetic distance estimate(s) of each suspected NUMT.....	83
Table 4.3 Results of referencing spurious sequences obtained from fecal samples in this study against <i>Tursiops truncatus</i> draft whole genome shotgun sequences on GenBank by the BLAST algorithm. Grey shading denotes homologous matches of $\geq 98\%$	85
Table 5.1 Summary of Sarasota Bay samples for telomere assays.....	92
Table 5.2 PCR assay parameters for q-PCR telomere assays.	98
Table 6.1 Summary of methods available to study the diet of live cetaceans.....	119

Acknowledgements

I thank my supervisors Nick Gales, Simon Jarman and Mark Hindell. Nick for finding a place for me in his research program, opening up doors that would usually remain closed to me and offering advice for developing my still appalling diplomatic skills. Thanks to Simon for offering jovial advice on many things molecular and science writing, and both Simon and Nick for facilitating my research through their institutional funding. Thanks to Mark for providing timely and insightful academic guidance throughout the duration of my candidature as well as a place in his lab at UTAS (which, importantly, has a fridge with beer in it most of the time).

Funding for this project was provided by the Australian Antarctic Division, the Winifred Violet Estate Trust Fund, the ANZ Holsworth Wildlife Research Endowment, The Norwegian Polar Institute, The University of Tasmania and the Australian Centre for Applied Marine Mammal Science (now the Australian Marine Mammal Centre). Further support was also provided by Mote Marine Laboratories and the Chicago Zoological Society. I am extremely grateful to these organisations for their financial support of my research.

I am grateful to Randy Wells, Nelio Barros (deceased), Dammon and Janet Gannon, Jason Allen, Aaron Barleycorn, Brian Balmer and all staff and volunteers of the Sarasota Dolphin Research Program. The work on this thesis on Sarasota dolphins would not have been possible without these people. I especially thank Randy Wells for being generous in his support and facilitation of my research. Thank you to Ruth Casper who performed the dolphin feeding trial that some of the work in this thesis was based upon and who allowed me to use the samples. Thanks to Sarah Laverick for helping Ruth with this feeding trial and also providing administrative support. Thanks to C. Scott Baker for help in facilitating the small Southern Right Whale component in this thesis. Thanks also to Christian Lydersen for facilitating my participation in the harp seal study.

Lastly thank you to my beautiful girl Krista for putting up with me (for the most part) throughout this process.

Abstract

Traditional molecular ecology has focused on describing the historic processes that lead to contemporary patterns of diversity within populations, species and higher taxa. Molecular tools can also identify the origin of biological material and shed light on contemporary ecological processes of populations and individuals. This thesis is concerned with evaluating the efficacy of some of these latter nascent applications for diagnosing near real-time ecological information in marine mammals. The applications under investigation were (i) DNA-based methods to identify the prey of cetaceans and (ii) using the size of telomeres within the life of an individual to estimate the age of individuals in the Pinnipedia and Cetacea.

Diet samples like faeces are complex mixtures of predator, prey and symbiont DNA and as such they require techniques that can exclusively target prey DNA. Previous DNA-based diet studies had employed species- or group-specific polymerase chain reaction (PCR) primers to achieve this and thus were implemented *ad hoc* or required *a priori* diet knowledge, which limited their scope. I developed a prey detection method that employed novel PCR primers widely complementary ('quasi-universal') for most animal 16S mitochondrial DNA (mtDNA) and a restriction enzyme to selectively exclude predator mtDNA. The method requires no *a priori* diet knowledge and can be applied to other predators with a minimum of modification. Faecal samples were collected from two sources; captive bottlenose dolphins (*Tursiops sp.*) fed a known diet and free-ranging bottlenose dolphins from Sarasota Bay, Florida.

Two techniques were applied to detect prey DNA in the captive samples; amplification of a small mtDNA fragment using a species-specific PCR primer pairs designed to detect a known prey species and the 'quasi-universal' method. Using the species-specific method, a prey signal was detected within 4-7 hours of feeding the captive dolphins the known diet and persisted for at least 12-19 hours after the diet ceased. After the first detection, 60 +/- 12% (mean +/- 95% CI) of captive samples contained a prey DNA signal using species-specific methods. The 'quasi-universal' method was applied to 12 samples from within the time period with a known diet of 10 prey species from 3 Phyla. Up to six prey species were detected per sample (range 0-6, mean 3.2 +/- 1.7 (SD) species) and all but one prey species consisting of 2% wet weight of the total diet were detected across all samples. No prey DNA was detected from one captive sample using this method. Estimates of prey item amplicon amounts showed congruence with the total proportion of wet weight of most prey items in the diet, though variability introduced through sampling amplicon clone libraries puts wide confidence intervals on these results.

The 'quasi universal' method was then applied to 15 faecal and 9 gastric samples from 19 free-ranging Sarasota Bay dolphins. Thirty two prey molecular operational taxonomic units (MOTUs) were identified across all samples (range 0-9, mean 3.7 +/- 2.2 per individual), consisting of 28 taxonomic assignments, 18 of which were species level identifications. One

sample did not contain prey DNA. The difference in results between samples from the captive animals and free-ranging animals suggest that factors such as sample collection methods, sample amount, sample storage duration and whether animals consume live or dead prey may affect the efficacy of DNA-based techniques, which has ramifications for interpreting results from captive feeding trials. These results were also congruent with diet data from this population via traditional hard-parts analysis of stomach contents from stranded individuals.

An unexpected consequence of using restriction enzymes to exclude predator mtDNA was the appearance of nuclear mitochondrial pseudogenes (NUMTs) in samples. The appearance of NUMTs in 15 faecal samples from Sarasota Bay dolphins was further investigated, in order to understand their impact on DNA-based dietary analysis in a field situation. Nine unique NUMT paralogs detected in 13 of 15 samples were represented by 1-5 paralogs per sample and were estimated to be between 5-100% of all amplicons produced per sample. The diversity of prey DNA and the proportion of NUMT amplicons per sample were related to real-time PCR cycling characteristics, with lower prey diversity and a higher proportion of NUMTs recovered with increasing real-time PCR threshold cycle values. This indicated that low DNA yields from diet samples are more likely to have NUMTs detected and less likely to contain prey DNA using this technique. This predator-prey system is relatively well sampled, which facilitated ease of identification of NUMTs, however for many study systems this may not be the case.

Telomeres are nucleoprotein structures on the end of eukaryote chromosomes that consist of regions containing 'telomere-like' and 'true' telomere tandemly repeated DNA sequence and single stranded telomere sequence overhangs on the 3' end of each anti-parallel DNA molecule, each with associated proteins. They change size throughout the life of many animals, suggesting that they be a molecular means to estimate animal age. To examine whether telomeres would be useful to estimate the age of pinnipeds and cetaceans, samples were collected from populations of three model species where the age of individuals was known or could be relatively inferred; harp seals (*Pagophilus groenlandicus*), bottlenose dolphins and southern right whales (*Eubalaena australis*).

In Harp Seals, telomeres were measured using two techniques, (i) de-naturing terminal restriction fragment analysis and (ii) quantitative PCR (Q-PCR). There was no relationship between age and telomere length using either telomere measurement method, however there was a strong correlation between the methods, indicating that they were comparable. Telomere dynamics in cetaceans were then investigated. Previous studies had shown that satellite DNA in Mysticete cetaceans contains telomere sequence repeats that may bias telomere measurement techniques. This was investigated in Odontocete cetaceans by characterizing interstitial telomere sequence (ITS – that is, telomere repeat DNA sequence that is not a part of true telomeres) in Bottlenose Dolphins. It was found that substantial ITS exists in bottlenose dolphins, and given its presence in closely related Mysticete cetaceans, most likely all cetaceans. The presence of this ITS made denaturing TRF analysis difficult to

interpret and so attempts were made to measure Bottlenose Dolphin telomeres using non-denaturing TRF assays and Q-PCR assays. Attempts at non-denaturing assays were unsuccessful and further efforts focused on Q-PCR assays. No relationship was found between Q-PCR telomere metrics and age in bottlenose dolphins. In four instances longitudinal samples were available from the same bottlenose dolphin individual a number of years apart and these were compared. While two samples showed the typical pattern of decline in telomere size with age, one showed no discernable change with age and one individual displayed an apparent gain in telomere sequence with age.

Given these results a small subsection of 6 cow-calf pair Southern Right Whale samples were initially analysed. Four adults appeared to contain shorter telomere sequences than their calves although in two cases, calves contained less telomere sequence than their presumed mother. In light of these results the use of telomeres to estimate age of individual marine mammals did not appear a valid technique and could not be recommended.

Overall, this study achieved its aims of appraising the efficacy of these nascent molecular ecology techniques in marine mammals. In the first instance, DNA-based diet analysis appears to hold great promise for analysis of cetacean diet. The methods were sensitive, identified prey to the lowest taxonomic level in many cases, and made use of samples that cannot be used for traditional diet analyses. They allow high-resolution prey detection of live animals, a feat that cannot be otherwise achieved in cetaceans failing direct observation of feeding events. Additionally the 'quasi-universal' method can be applied to any predator where its 16S mtDNA sequence is known or can be gained, and the biases of the technique can be inferred by using current data from databases such as GenBank. Conversely, telomeres appear to hold little use for age estimation in marine mammals. All methodological issues aside, there are many exogenous and endogenous influences other than chronological time on individual telomere dynamics and these are not well understood. It is not recommended that telomeres be used for age estimation in any animal group without considerably more work. The latter outcome is as useful as the former, since any emerging technique (no matter how promising) must be put through rigorous critical appraisal in order to understand whether the application is warranted at all, or if so what the caveats might be.